D4.1.2: MATREX MODIFICATION

Ankit K. Srivastava, Declan Groves

Distribution: Public

MultilingualWeb-LT (LT-Web)
Language Technology in the Web
FP7-ICT-2011-7

Project no: 287815
## Document Information

<table>
<thead>
<tr>
<th>Deliverable number:</th>
<th>4.1.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable title:</td>
<td>MaTrEx Modification</td>
</tr>
<tr>
<td>Dissemination level:</td>
<td>PU</td>
</tr>
<tr>
<td>Contractual date of delivery:</td>
<td>31st January 2013</td>
</tr>
<tr>
<td>Actual date of delivery:</td>
<td>15th February 2013</td>
</tr>
<tr>
<td>Author(s):</td>
<td>Ankit K. Srivastava, Declan Groves</td>
</tr>
<tr>
<td>Participants:</td>
<td>DCU</td>
</tr>
<tr>
<td>Internal Reviewer:</td>
<td>DCU</td>
</tr>
<tr>
<td>Workpackage:</td>
<td>WP4</td>
</tr>
<tr>
<td>Task Responsible:</td>
<td>Declan Groves</td>
</tr>
<tr>
<td>Workpackage Leader:</td>
<td>Pedro L. Diez Orzas</td>
</tr>
</tbody>
</table>

## Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Author</th>
<th>Organization</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13/02/2013</td>
<td>Ankit K. Srivastava, Declan Groves</td>
<td>DCU</td>
<td>Final version</td>
</tr>
</tbody>
</table>
CONTENTS

Document Information........................................................................................................................................2
Revision History..................................................................................................................................................2
Contents ..........................................................................................................................................................3
1. MaTrEx Modification ..................................................................................................................................4
   1.1 Link to prototype or downloadable ...................................................................................................4
   1.2 Description .........................................................................................................................................4
   1.3 User guide ...........................................................................................................................................4
1. **MaTrEx Modification**

1.1 **Link to prototype or downloadable**

http://www.cngl.ie/mlwlt/

1.2 **Description**

The Dublin City University (DCU) Machine Translation (MT) system, also known as MaTrEx (Machine Translation using Examples) is a statistical MT system using the open source log-linear phrase-based decoder Moses (http://www.statmt.org/moses).

The MaTrEx engine for MLW-LT can translate in 4 language directions:

- English \(\rightarrow\) French
- English \(\rightarrow\) Spanish
- French \(\rightarrow\) English
- Spanish \(\rightarrow\) English

It can take as input text in the following formats:

- plain-text (one sentence at a time)
- webpages (HTML5)
- XML documents

The current MaTrEx configuration can handle the following ITS 2.0 data categories:

- Translate
- Language Information
- Domain
- Locale Filter
- MT Confidence

1.3 **User guide**

The DCU MT web service for MLW-LT has been implemented using the Soaplab web services software framework (http://soaplab.sourceforge.net/soaplab2/). It provides a RESTful web service. The MaTrEx web service can also be accessed directly by logging on to the website as shown below.
Scroll down and click on “translate_main_dev” link. This will open the interface for translating as shown below.

To translate, either enter the webpage to be translated (URL, input field 1) or type in the text to be translated (input field 2)

Select the source language: en (English) | fr (French) | es (Spanish)
Select the target language: en (English) | fr (French) | es (Spanish)

Currently, the MT system can translate in the following directions: en↔fr and en↔es

Click on “Run Service”. The output will be displayed in a separate screen.